So just how does a brain work and can we design machines the same way?

Professor Keith Kendrick

26 May 2005

We are going to consider between us, to some extent at least, how the brain actually works, or how it might be working, and could we possibly in the next 50 to 100 years really start to design machines to work the same way?

Unfortunately, for anybody who has worked on the brain, compared to other parts of the body, what seems to provide a reasonable understanding for the way your kidney works, or your heart works, or your liver works, or testes or something like that work, really doesn't give you even first base understanding of how a brain works. We really don't have an idea about how it allows us, and of course other species with brains, to have some kind of complex experience of the world around us, and for that matter, even experience themselves.

I'm not alone in asking these kind of questions, as you might imagine. In fact, about 25,000 of us go for an annual trip to the United States where the Society for Neuroscience meets. I estimate that probably – this may be an underestimate – there are at least 50,000 of us working around the world trying to have some idea about how the brain actually works, although I would say that for the vast majority of these individuals, any kind of big question like how the brain works is not something that's on the agenda. Everyone breaks the problem down into much, much tinier components. It's only those of us, when we get a little bit older, who have done all that and decided to step back, who say, well okay, I've spent all this time trying to understand the little micro-component and it's not really told me anything about how the whole thing works. So only when you get a bit older do you get brave enough to give talks like this, and probably go down in flames as a result of it!

So if I, or any of my colleagues, could explain to you how the brain works, what would it give us? Well obviously it would help us very much in understanding and being able to treat the whole host of brain dysfunctions that we are subject to: developmental ones like autism, Aspergers; psychiatric ones like depression, schizophrenia, and of course neuro-degenerative ones, either just simply senile dementia, or Alzheimer's or Parkinson's. Really, we don't have enough understanding of the way the brain works to be able to treat successfully any of these problems.

Of course, those of you who have listened to the majority of my other lectures will know I spent a lot of time talking about what other animals are capable of compared to ourselves. If we could understand how brains work in principle, then we should be able to have a much better idea about what the real mental capacities are of other species, as well as of course our own.

For those of you who have been itching for probably the last 10 or 20 years or longer to get rid of that wretched computer keyboard and to be able to interface directly with your computer without any of this is sort of mechanical and slow stuff, it could allow us to directly interface with machines – a brain/machine interface.

And of course finally, the real science fiction, only in Hollywood at the moment, the idea is that we would be able to perhaps construct bio-robots that were as good as we are, but possibly even better.

Let's step back for a moment and remember how remarkable the brain really is compared to the other technologies that we boast about and accept as part of our every day lives, particularly obviously computers. The brain is about 920 cubic centimetres. It's about one and a half kilos. If it was much bigger than that, it would be very hard for us to carry it around on our shoulders. It has a staggering 100 billion nerve cells within it, and because each of those nerve cells can form 1,000 to 10,000 connections with other nerve cells, you end up with a very big number, one quadrillion synaptic connections. If we work out that synapses can transmit information at maybe about 10 bits per second, you end up with a sort of global information processing capacity of a brain of 10 quadrillion synapse operations per second.

Not surprisingly, all that electrical activity burns an awful lot of energy, so your brain, hopefully when you're listening to me, is burning off at least 10 watts. It might be a rather dim light bulb, but that's an awful lot for a body to generate. The brain is energetically the most hungry organ in the body.

So what can our computer experts come up with? The biggest one – Aski-Purple, created by IBM, at the US Department of Energy is 830 square metres in size. It weighs in at a massive 200 tons. And it still only has half of the information processing capacity of a brain…

Of course, no one would argue for a moment that the brain does everything as well as computers. In fact, there are a lot of things it doesn't do anything like as well as a computer. It's trying to do rather different things. One of the things that brains do compared to computers that makes them so clever is that they carry out a lot of interpretation of what we think should be going on, even if it isn't actually going on at all. One of the key areas that all brains have is the ability to attend selectively to something going on in the environment and pretty much ignore everything else.

I will now play a film of students playing basketball, and your job, playing with your selective attention, is simply to count how many times they pass to one another. Just to make life more difficult, there are 2 basketballs. They're fairly slow, I haven't made them too fast, but there's 2 balls, and you've got to count how many times they actually throw them to each other.

So how many times did they do it? Did anyone notice that there was something else going on?

We'll run it again. Now don't bother to count the balls. Just watch – there's something unusual happening in the sequence – and here it comes! Someone in a gorilla suit comes on and basically makes fun of you!

The vast majority of people seeing that for the first time, given the instruction to count the number of times the basketballs are passed, will not see the gorilla at all! It's a great example of how selective attention works as far as the brain is concerned.

Another very simple thing that all brains do, in the visual system at least, is to ensure you can work out your optics. The lens of your eye actually turns things upside-down, but we all know that we do not see the world upside down and nor, for that matter, do we imagine that other species with eyes see things upside-down. The brain is what turns the world back the right way up. It assumes that's the way it appears, so that's the way you experience it, but it is the brain interpreting something that's upside-down.

Sometimes your brain thinks things should appear the right way, so it interprets it that way, whereas in fact, it looks anything but normal.

So the brain is a phenomenal interpreting device and one should never forget that. It has very big ramifications, particularly when we're talking about, for example, our witness-based judicial system. I'm sure you all remember things, particularly that happened to you early in life, these fond rose-tinted memories. Life was fantastic when you were young, everything was positive and great and all the rest of it, and of course you believe that what you remember was the way it actually happened, but it might not be, because every time the brain remembers something, it rewrites it, and it will rewrite it in the context of whatever's going on at that particular point in time. If you've had a bad day, or another experience, they can get mixed up, and so when you lay down the memory again, it's no longer exactly as accurate as it was when the event first occurred to you.

This was shown experimentally in a study carried out in the United States for people trying to remember the facts of the OJ Simpson trial, which may not have been very big in the UK, but it certainly was in the United States. When people were asked to recall the facts, in terms of accuracy and distortions, the majority of people got everything right and there were very few factual distortions after 15 months. However, at 32 months, there was a big drop in the accuracy and a large increase in the number of distortions. Given how long it takes crimes to go to trial, these kinds of periods are fairly relevant, so people may not remember very accurately what happened when they act as witnesses. But certainly, the brain does have this problem, unlike a computer: once a computer has stored something, that's it. The brain is constantly updating it and changing it every time you remember something.

The thing that everybody likes to talk about and nobody has a really good answer for, but what distinguishes brains, or certainly human brains, from computers, is consciousness. We know perfectly well that computers are not conscious. We know less really as to whether the beetle, which has a brain-like structure, is capable of consciousness, or whether the cat, or the chimpanzee – all we know is that we are. Probably most people would argue that it is an inherent property of brains, consciousness, although you have to achieve some degree of complexity, and possibly the algorithms that are used to link up the networks in the brain, in order for some level of conscious experience to emerge.

So if we're considering how a brain might work, of course you can look at it at a large number of different levels, all the way from genes through to the molecules of the brain and how they are released, for example, from synapses, how little networks of cells group together to perform functions, and ultimately, how when we put them altogether into the brain, how do they work in some kind of holistic function to give more than the sum of its parts, because believe you me, that is the secret of the brain. It isn't that you just add all the parts together and you get function. The brain is doing something, or lots of things, holistically, and that's one of the main things that distinguishes it from a computer, which does not.

We know how important the brain is. Two-thirds of our 30,000 or so genes are actually expressed in the brain. It's by far the greatest level of expression of our genome in any organ. We really need to understand how brains work as a cohesive, communicating network of cells, and that's a pretty mind-boggling task given how many billions of nerve cells there are. However, one advantage is that they use quite a simple language, but then again, the language I'm speaking to you is pretty simple too; it just has infinite possibilities when you put all the combinations together, and so it is the same with brains. A simple language can be converted into an infinite number of possibilities.

As far as the highest level of communication is concerned in the brain, primarily that communication is simple, electrical impulses. If we convert it back to language, at any particular nerve cell has many thousands of cells talking to it, saying either yes because they're excitatory so they use a particular kind of chemical transmitter, or they're saying no, don't do anything, they are inhibitory input and they use a different transmitter. When there's a balance between the "no"s and the "yes's", it doesn't do anything. It all cancels out. It doesn't send any messages off downstream. However, when the "yes's" in this particular case are occurring more frequently and strongly than the "no"s, then this fires off a "yes" message from the cell to its target down here, and then that starts the whole process on downstream. Very importantly in getting the complexity of the brain, there's feedback, so it's also telling the guys that are signalling to it and saying things to it, "Hey, I'm saying yes," so there's feedback going on as well, and this very simple system is capable of huge complexity.

Just to make life a little bit more difficult, we get conduction of electrical impulses from the cell body through the axons. It's one way traffic as far as the nervous system is concerned, it's only transmitting one way, although in fact the axons can transmit the other way artificially. When they get to the end of the road, there's a gap between their process and their receiving cell. It's a very small gap, it's less than a micron, and the information is transmitted across that gap chemically. So there's a release of chemicals from the pre-synaptic terminal at the end of the axon. These pass across the synaptic cleft, and act on various receptors to cause changes in the electrical activity of the target cell, either depolarising it or hyperpolarising it, and so propagating the message on to the next cell. This is of course capable of infinite control.

All these simple things that I've so far described to you have already gathered a clutch of Nobel Prizes. They were very difficult to discover just at this level. So you've got recognition of the structure of the nervous system, Golgi and Cajal; you've got functions of neurons, Sherrington for example, and Lord Adrian; and for the chemical transmission of nerve impulses, synapses, Sir Henry Dale and Otto Loewi. There are of course many other individuals who have received prizes for their work on basic elements of functions of nerve cells. But at this stage, we know how they work. We just don't know what it is they're saying, that gives us experience. So what are the messages really saying? It's a kind of huge enigma machine that's in there that is somehow or other conveying huge amounts of complex information and it's also giving us a conscious experience. It's not just what they're saying, it's also who's listening to them how are they interpreted? And finally of course, the real $64,000 question is, how do they give rise, all this electrical activity, to our perception, thought and action, things we take for granted?

The first obvious thing when you start looking at brain function is that it is modular. It isn't some kind of homogenous structure where everything is capable of doing everything else. If you start looking at it in any detail, you can see their lumps and bumps and nuclei and so forth. The phrenologists thought they had it all worked out as to exactly what different parts of the brain underlying the skull were doing. In a sense, they have it right, but in another sense, they have it completely wrong. It certainly is modular.

Wilder Penfield, who was a surgeon, when doing surgery on humans would electrically stimulate parts of the brain to see what they were doing before he operated on them. The areas he looked at were areas like the motor cortex, so if he stimulated a tiny bit of the brain, the hand moved, or a finger moved, or a foot moved, and you could see a very, very clear concordance between a tiny amount of electrical stimulation to an area of the brain and some kind of behavioural outcome.

We also know, unfortunately, both from accidents but also from experimental work, that sometimes even quite tiny bits of damage to this complex organ can cause very selective problems in terms of our behaviours. So for example, if you have your visual cortex destroyed, you won't be able to see anything, or at least you don't have any experience of seeing. If you have other parts of your brain destroyed, you might not be able to remember anything, and so forth. There are a number of parts of the brain, like for example language on the left hand side, Broca's Area; if that's destroyed, you won't be able to speak. These are very obvious examples of what appears to be modularity of localised function.

Functional magnetic resonance imaging studies have a control situation where the subjects go into the scanner and that gives a background, and then you have some kind of stimulus that they see, a word or something like that, and you get another pattern of activation and you subtract one from the other, and what you see when they represent their data is the difference. It doesn't mean this is the way the brain works, but what you do get is this highly localised pattern of activation that appears to be associated with a particular stimulus. That doesn't mean that's the way the brain works, but all this is telling you is that if you compute the difference between what the brain is doing before and during a particular stimulus, this is what you get. In fact, the majority of the brain is active during processing, both during the control time and during the stimulus time, but it does give you this idea that there's a modularity.

If we look at the senses, you get exactly the same modularity impression. They are all beautifully divided up, the five senses. We have touch, we have taste, we have smell, we have vision, and we have hearing, all in physically different parts of the cortex. The systems are kept very much separate.

Even elements of identifying things as what they are, as opposed to where they are, are treated by completely different systems within the brain. All of these different systems are capable independently it appears of consciousness, so there is no single seat of consciousness. Descartes thought it might have been pineal gland on the simple argument that it was the only structure in the brain that had a single representation. It wasn't bilateral, and you can't possibly have two consciousnesses, so that must be where consciousness resides. In fact the pineal is very important for producing melatonin and dealing with circadian rhythms, but it certainly doesn't have anything to do with consciousness. Indeed, there doesn't seem to be a single area of the brain that allows you to be conscious. All of these systems are capable of providing you, independently, with a consciousness of that particular modality but of course we don't experience six or five or whatever different consciousnesses simultaneously. We only experience one. Something that the brain is doing is linking up all of these separate systems so we get a unitary experience rather than five or more different ones, but in reality, there are five or more different ones, as far as the brain is concerned.

You're probably getting sceptical about this, wondering why we should have our sensory systems so delineated in this way. The reason can be really shown quite clearly by individuals who suffer from a condition called synaesthesia, where they tend to mix their sensory modalities, unavoidably. For example, primarily, it's seeing or hearing words or letters as colours. Just an example of this one, a person actually, with words or names, experienced tastes. In one case, a potential girlfriend called Tracy evoked a taste of flaky pastry, which you can imagine was likely to kill the romance rather quickly! As I say, these are unavoidable. It's not some sort of learning thing, it just happens – you can't stop them coming into consciousness. There are other individuals who effectively taste shapes, some who literally feel as a physical sensation musical instruments, or the sound of musical instruments, in specific parts of their bodies. It affects something between one and 201 in 20,000 people, and, not surprisingly, it's very prevalent in artists. One study reported as many as 23% of fine arts students exhibiting some degree of synaesthesia.

It's thought to be caused exactly by the problem that normally the brain sorts out by keeping everything separate. That is, it's probably due to cross-wiring between the adjacent sensory maps in the brain. Indeed, Simon Baron-Cohen, for example, in Cambridge, has argued that it's probably something that we all have as infants, that everything's tied up in this way, but we grow out of it, all the connections between the different sensory modalities die off during the early years of our life, so that by the time we get to full maturity, there isn't any tie-up between the different modalities. You need to keep them separately if you're going to be able to unify your different states of consciousness with the different systems in a way that doesn't get them mixed up.

So, we've got the different modules, we know they're there and they're all capable of independent function, but how do they become unified? First of all, if we look at how the modules function, they are beautifully organised into special structures, where different parts are dealing with different components. For example, in motor cortex, or somatic sensory cortex, the feet, the body, the hands, the fingers, they're all dealt with by a physically separate part, a different part, within the module.

If you look at pretty much any of the sensory systems, in fact, as far as I know, all of the sensory systems, it's quite easy to show that what is going on in terms of the stimulus in the outside world is represented fairly faithfully in a special pattern within the brain. But you have to be very careful in taking the next step from this, and the idea that somehow or other, the world is represented as a spatial pattern of activation in your brain, and that there's a downstream interpreter, a little man in the brain, looking at the pattern, and saying, "Hey, that's a wheel," or "that's something else" because you're immediately into a dualistic fallacy which will go on in perpetuity if you want to. You cannot have a concept where these patterns are being interpreted independently by a separate part of the brain, because you immediately fall into this dualistic brain-mind type division. Unfortunately, no matter how you look at it, you have to consider that the activation patterns you see may be spatially organised like that, but they are intrinsically what you experience. It's not a two-faced system. It's a single process. What is activated in the brain is also giving rise to your experience of that thing.

If we look in a little bit more detail about the way that neural networks function within these modules, you start to see perhaps some principles of the way that information is encoded. It's still a bit theoretical at this stage, but nevertheless, the idea is, it's almost like an inverted triangle, you start off with lots of the network dedicated to very simple patterns, like lines – we're now talking about face recognition, or how a face recognition system might work. So you have a large number of cells just dedicated to detecting lines of a particular orientation. You have equally quite a large number of cells responding to colours. You have another large number of cells trying to understand the outline of a face – that's the right kind of egg-shaped pattern. You have another set of detectors – you're getting smaller and smaller numbers all the time – that are looking at the features – eyes, nose, mouth and so forth. And then we start to get into the more holistic, processing regions. A smaller number of cells deal with: "Is that a face as opposed to another kind of object?" "Is that a particular type of face?" European, Japanese, etc., and down even further, "Is that a face that I know?" and finally, maybe even down to the point of "That's a particular individual."

We know that these kinds of cells can be found in the human brain. It can be shown they exist even in the brains of sheep, though they seem to be tuned just to a single individual's face. However, it's a fallacy to consider that our ability, or a sheep's ability, to recognise a face is totally dependent on one cell, that there may be only several hundred in the whole brain that will respond in this way. This is not the point at which, when this is activated, you say, "Ha! That's my grandmother. That's my mother." The actual recognition is a property of the whole system, not just the final end point, and indeed you can still get the impression of your grandmother without all of the whole network being activated.

But one important thing about organising a recognition system like this is that by building up levels of hierarchy, you can start to reduce the numbers of the components that are involved at the different states, so that you end up by increasing the number of stages that you can process information, but the amount of the neural network that's contributing to each of the individual stages becomes less and less. So learning something, or making something familiar, ends up by making a sparsening of the system. It has less of it dedicated to performing the recognition function than it did at the start. Of course that makes sense. What you don't want is to have more and more and more of your brain trying to recognise something that you've learned about. You want less so you can use the brain for other things. This is quite an important principle, and it does actually operate within the brain.

But how do all these different levels link up? For a long time, it's been felt that perhaps one of the important principles of the way that these vast networks of cells link up is some kind of synchronisation process, a binding process. This is just showing theoretically what might be going on, and in fact, it's probably not what's going on, but nevertheless, this is what most people thought was going on.

The idea is, within an unfamiliar face, you get cells responding to non-specific aspects, so not just a face but components of faces. They're not firing synchronously, they're out of synchrony. They may in turn activate cells that are saying, "Well, that's a particular kind of face," and again, they're not synchronous and you don't get translation of their activity through into higher order processing.

Whereas if we look at a familiar face, then we get synchronisation going across the network, right from the early processing through to the "Is that a face?" through to "That's a particular kind of face". What we would argue is that you need less and less of these cells firing in this synchronous manner to be able to link up these three different levels of processing so that you get an accurate form of recognition. So the synchronisation is binding together information from the different levels, and it's also making the information processing more efficient.

If we look at learning in general, a Canadian psychologist, Donald Hebb, came up with probably the simplest and still very influential ideas about how learning influences the cell in a nervous system. It's very simple: literally, when an axon of a cell A is near enough to excite a cell B and does so repeatedly or persistently in firing it, then some growth process or metabolic change takes place, in one or both cells, such that A's efficiency in firing B is increased. It's about as simple as you can get, but that seems to be the way that the nervous system works. Not surprisingly, if A leads to B repeatedly, then you need less and less of A to get B to do something, and so it gets more efficient, but you also tend to get more synchrony between what A is doing and what B is doing.

Stepping back away from single nerve cells, if we look at the global activity of the brain, we find it somewhat puzzling, and it will be interesting when we come back later in the talk to why this may be relevant, but our waking brainwaves from the cortex are high frequency, low voltage and they are desynchronised. They are not doing things at the same time. However, as you, for example, lie down to sleep, they slow down and start – they're still desynchronised at this stage – and then finally when you start to go into the early stages of sleep, a dramatic change occurs. They become beautiful rolling synchronised slow frequency waves, which are called theta. But it isn't just sleep that causes these synchronised waves across the cortex, it's also things like learning, which has led to this idea that somehow or other, when we learn things, there's some kind of global synchronisation process going on in the brain that allows it to operate in a more efficient, in a holistic way, so that information is tied up between one part of the brain and another.

Almost everybody's had a go at trying to explain how consciousness occurs. Susan Greenfield said that perhaps it's a bit like a stone being thrown into a pond, and gradually the ripples emanate out from that stone, and affect more and more and more of the brain. That may be true, but whether that's the process by which you become conscious, is a question, because it would then imply that it's like a mass action thing, that you have to have a huge amount of activation of the brain, and suddenly you become conscious when a large amount of the brain has become activated after a small part has processed a particular signal.

That's one way of looking at it, it's quite simplistic. The way that I think is more important, given the modularity of the brain, and they're all doing their things separately, but somehow or other, they're all joined up, they give us a unified experience of consciousness, you have to come and think about like different elements of a picture. Suddenly, in themselves, they don't evoke any consciousness at all, but when they are built together and they overlap and they weld together, that is the point at which you are conscious, and all those things are being dealt with by different parts of the brain. Something has to link them together, to bind them, so that suddenly, whatever it is you're looking at, instead of just being unconsciously processed, you are fully aware of in consciousness. It's like the idea of four different pictures converting into a single scene. So in different attributes of an object, the brain is managing through coordinating activity in some way, to bring together, as a whole, and when it does that, that's the point at which you're conscious of it. This idea has been put forward by a number of people, but notably by one of the co-discoverers of DNA, the late Francis Crick, who also showed an interest in consciousness in the brain in the latter stages of his career, even though he wasn't a neuroscientist.

Experiments have tried to look at what it is that changes from the moment that something is happening to you and you're not aware of it to the moment suddenly you are aware of it. In order to manipulate these kind of things artificially, you need to take something fairly simple that can be manipulated.

Derek Denton at the Howard Florey Institute in Australia, and other groups, have tried to do this using two very simple systems. One is hunger for air. As you starve people of oxygen, if you do it very, very slowly, you're not aware of your need for gulping air, and then suddenly, it hits you like an express train – you need air desperately and you're aware of the problem. Equally, thirst is another one. You can put salt into people's bodies and there will come a point at which you will develop a raging consciousness for thirst.

It kind of illustrates two points. It's really not the case that something uniquely becomes activated when the individuals become conscious of their need for water or their need for oxygen. It's a gradual build-up in the activation, particularly of cortical areas that control attention, like the cingulated cortex. It does not however mean that these areas are essential for consciousness. It's merely turning the gain up in the system, and particularly in the near cortex which I think most people would imagine is going to be the most important part of the brain for allowing the experience of some kind of consciousness.

These are broad principles we're talking about here, but how do these large networks really work, and how do we find it out? There's only a simple way of doing it, but it's extraordinarily difficult to achieve, and that's to listen in on their conversations, and not just one or two of them, but hundreds of them while they're doing something. That technology has not really been available to us, until recently, and so the how is this: it's using lots of bugging devices, and this is the area of research that I currently spend most of my time with, and it's giving us some very useful insights now into how brain systems actually work as networks.

These are like very, very fine needles on a grid. You can have 100 of these sort of bugging devices that you can put into a few millimetres of cortex, and these will allow you to record the electrical output of perhaps several hundred nerve cells at the same time, while the individual is exposed to a particular kind of stimulus, and in our case, we deal both with smell and also with faces.

With the smell system, it's a relatively simple one and we know a lot about its organisation, so it's a very good place to start. You have olfactory receptors or odorant receptors in the nasal epithelium. There are many thousands of them, and they send the signals from their chemical stimulation to the olfactory bulb, which does an awful lot of processing of smell, and then sends the result of the integration off to other parts of the brain that are controlling behaviour.

The olfactory bulb is a fairly simple lamina structure which allows you to put electrodes just into looking at one row or one particular part of the processing network at a time. The deep levels of processing in the olfactory bulb is where most of the clever work is going on in terms of teasing out the difference between one odour and another.

Just to give you an idea of how this system works, you have five million sensory neurons picking up the odours in the epithelium, and they have a staggering thousand different receptor types. The olfactory bulbs though, each one of them has 2.5 thousand pick-up points, where they're taking in all this information, so you can see the whole thing is coming down at a phenomenal rate. Remarkably, and no computer could ever do this with 2,500 transistors, it can distinguish 10,000 different odorants.

If we look at cells that just turn up or turn down their game, they're producing more or less electrical impulses, and that in the past has been all that we've been concentrating on, and that, for example, is all that an MRI scan will give you, whether something is producing more or less activity.

Well, the first thing we find is exactly what has been predicted from the idea that as you learn in the system, fewer and fewer of these cells change their activation patterns, that there is sparsening going on. So every time a smell is smelled, the system learns something about it, and less and less of the network needs to be dedicated towards responding to that cell.

But what about the other 90% of the network we're recording from? Are they just sitting there waiting for something that happens to be important for them to occur? The answer is, no, they're not. We're starting to reveal the principles of the way the brain works that are very different from the way computers work, and it's also beginning to show us why brains are capable of such phenomenal information processing.

If you record from a nerve cell, it isn't regular like a metronome, it's all over the place. It's like Morse code. Now you might think, okay it's just random – but it isn't. These are non-random patterns that are being produced within nerve cells. You can see them being repeated within the train of activity that you get.

In recordings from an olfactory bulb, if you get a fair amount of the olfactory bulb being recorded from, it isn't just that you get patterns within one cell repeated, once, again, again, again; you actually get synchronised, or if you like, you get patterns across the network, which involve many, many different neurons, different cells. You get a sequence going on again and again, where the pattern in one cell is then followed by a pattern in another cell, followed by a pattern in another cell. We can start to pick up a level of complexity, not only within the patterns produced by a single neuron, but in the sequences of patterns that are being produced by whole networks of neurons. You don't have to be an information theorist to understand that in terms of the amount of information, in bits, or how you want to express it, that are represented by these patterns both within cells and across groups of cells, this is far greater because it's involving a massive network. The tiny amount of information contained in that little population of cells simply turns the game up and down. So there's a huge amount of information processing going on in the network without changing anything in the game. They are not going faster, they are not going slower, they're just saying something different.

Of course we want to know what these things mean, and unfortunately, it's going to take us some time to try to decipher how important the patterns are functionally, but we do know they are important functionally.

If we look at what's going on when the animal smells something, you'll notice that during inhalation phase, there's a large increase when they have a stimulation from an odour in the number of patterns that you see. Not only that, the complexity of the patterns increases as well. So we have got lots of information, additional information, being processed by the whole system, not just by the few cells that change their firing rate up and down, and you can even see that depending upon the concentration of the smell you put in, you get a progressive increase in the length of the patterns that you see. So the patterns are encoding all sorts of important information, and future research will have to try to work out functionally which particular pattern means what particular functional outcome. But at least we're beginning to understand what is important in the way the output of the whole network is using not just the tiny bit that we've been able to look at in the past.

There is another area, and this perhaps may be a little harder to understand, but its application hopefully will be very easy to understand.

Now, I said earlier that we all expected, and indeed most people only expect to find synchrony in the brain, the idea that brain systems somehow or other bind information by doing everything together. After all, teamwork is very much at the heart of our own ethos about how to achieve things. Doing things together is the way that we know we can actually solve problems. But it appears that the brain doesn't use that principle. In fact, what it seems to be using is not doing things quite together. So instead of I push, you push, it's I push, you pull. It may seem like a strange thing, and it's only a tiny shift. You could have all sorts of possible relationships between all of the cells that are firing away quite merrily in a network, and this just plots in pseudo-colour what the relationship is between each cell. One fires, do they fire at the same time as another cell, or do they fire at a different time from another cell? And all you need to see from this is this is what happens; that you, the majority of the cells are not positively correlated with one another, i.e. they're not synchronised, they're actually primarily slightly desynchronised. When the odour is applied to the system, they become even more desynchronised, so they're not doing things together, they're doing things more and more differently. As I say, the more times that the system learns, the more negative the correlation gets driven. And you say, well yeah, so what? I mean why? What's so important?

You have to remember that the nervous system appears to have a drawback in that it has a lot of noise in it. Cells have to be active all the time because if they're not, they'll die, and so you end up with a huge amount of noise in the nervous system, and yet, this system is capable of distinguishing signals from noise very, very efficiently, but in fact, what it's doing is it's using the noise in the system to actually cancel out the noise that is introduced into the system artificially with signals. It's doing it through having the network negatively correlated. So if the cells were not actually relating to one another at all, they were random, stochastic, whatever you want to call it. If, in a computer simulation, you have the sound wave in green, you can see that there's associated with hundreds of repetitions an element of noise, high frequency blue stuff. Now if we decided just to tweak the system and make it slightly positively correlated, only plus 0.1 you increase the noise. In fact, this is mathematically absolutely predictable from central limit theorem, which is the basis of estimating variance in statistics. It's exactly what happens. You amplify the noise by positively correlating it. However, if you negatively correlate it, or look at level you wipe out the noise entirely. This is what we believe is one of the most efficient systems that a noisy brain has for wiping out noise and making signals very, very easy to detect.

The other thing that having desynchronisation in the system gives you is a little bit more complicated to understand, but it explains why brains can do things in such a small size compared to computers. It's also to do with negative correlation. The only way I can try to describe it to you is that when you have a negatively correlated network, whatever it happens to be, and in comes a pattern of information, when it activates that network, because it's negatively correlated, it has the effect of like exploding that pattern in a theoretical physical space, so that it occupies pretty much the whole of that physical space. Whereas if it's positively correlated, it doesn't do that. The pattern is represented in a much smaller part of the theoretical space. If you're trying to discriminate one pattern, or any pattern, we all know that it's much easier to discriminate a pattern if the elements of that pattern are spatially further apart from one another. The further apart they are, the easier it is for us to say, that's one pattern, and to be able to discriminate it from another pattern. So by having a negatively correlated system, the brain can not only get rid of noise, it can also deal with overlapping patterns. It doesn't cause it any problem at all in distinguishing one pattern from another, and therefore it can use, like in the smell system, only 2,500 cells to pick out 10,000 different patterns. A computer can't do that.

It's not just smell. This also happens with the face system, which is association cortex, a much more sophisticated part of the brain. You also get sparsening of the small number of cells that respond to faces as the animal learns about faces, and you also get a decorrelation shift, even in this advanced part of the brain. So it's not just something to do with smell; it seems to be, as far as we can tell, a universal property of neural networks, at least in mammals.

And just to even show another element of this, we know that, for example, the right and the left parts of the brain for face recognition seem to do rather different things. Recognition is particularly involving the right side rather than the left. If you look at the levels of desynchronisation that are occurring when the animal is viewing faces, between the left and right, when they don't really know what they're doing, they don't know which face is which to get them the food reward, you can see that the overall levels of correlation are very different in the left and the right side of the brain. However, when they learn, they match up, so the organisation on the left and the right have become synchronised. They have the same level of correlation, which presumably is making them more efficient, and possibly this is also to do perhaps with the emergence of a consciousness as the result of the two sides pretty much being organised the same way at the same time.

Indeed, it isn't just learning that reveals this difference between the two sides. If the animal does know what it's doing but it gets it wrong, okay – normally when it gets it right, there's no difference between the two sides, but when it gets it wrong, there's a massive difference between the two sides. So either it's not conscious, or this is just a result of making a mistake. Also, you tend to get far more cells responding to the faces than you should do, so the system clearly is predictable in terms of whether or not it's not working properly in terms of what the two sides of the brain are doing in this one parameter of correlation.

So okay, you think well, this is the brain, and can we actually transfer this kind of approach to artificial systems. It took us a long time to work this out, because the first thing you have to do is to make signals that are captured by any kind of artificial device, some kind of sensor device, whether it's your video camera, or your hearing aid, or whatever, negatively correlated, in order to reduce noise and also in order to make the signals more discriminable. We found in the end that there's a very simple way of doing that – I won't go into detail, but effectively, it's taking a signal, a wave form, when you've captured it – it can be anything, analogue, digital – and shifting it by half a period, and adding the two, the captured signal to the phase transition signal, and that effectively negatively correlates the whole signal which you've got. It's a very, very simple thing to do, which is why we patented it.

Does it make a difference? Well, if you take a pixellated image with a huge amount of noise in it and we positively correlate the signals, we can see we make it an awful lot worse. If we apply the principle and we use a negative correlation, we get rid of the noise completely.

Similarly, we can do some important things with signals that may have relevance to telecommunications as well as, spy technology. This is called the cocktail party effect. If you remember, what I'm saying is that by having negatively correlated signals, you can actually make signals more discriminable, and one of the problems that we have, whether you're trying to transmit information or interpret it, is when you have mixed signals, all together. How do you bring out one signal from another? You can do this with a negatively correlated signal by applying a negative correlated component analysis. This is called the cocktail party effect – whether you're going to hear it or not. When four people are speaking simultaneously, we have an awful lot of difficulty understanding individual conversations. If we apply this negative correlation rhythm, we can pull out each one of the individual. So it's great for spies. You can pull out anything you want from a mixture of signals, and it means you could also send a mixture of signals down the telecommunications cable and decode them at the other end. You can afford to send them all mixed up rather than singly, so it's got all sorts of advantages.

The brain is able to teach us some amazing things – that we can already apply to existing technology to make it a huge degree more efficient, or at least that's what we hope. But how far are we going to go? Are we really going to be able to generate completely different computers and bio-robots? I'm sure most of you are aware of Moore's Law, the idea that a number of transistors that you need to have on a chip, a central processor unit, whether it's Intel or anyone else, but it's usually Intel, is required to double every 18 months. That's Moore 's law. If we get up to the current 2004, your Intel 4+, or whatever it will be, has 125 million transistors on its CPU. That's not enough. We can't make them any smaller, and so this is why Moore 's Law is regarded as being a big potential buffer at the end of the track, and people are looking at other ways of trying to create, in this case, organic switches rather than silicon based transistors. This is either quantum computing, if you're using atomic switches, or more recently, it's been suggested to use DNA, because you have a phenomenal amount of processing capacity just within the simple CGs and As and Ts of your DNA code. It was I think predicted that one handful of DNA could deal with most of the world's computing power all on its own.

However, even if we do this, none of these systems – they may be bigger, they may be better, they may be more efficient – I would argue that under no circumstances will they be conscious because they don't operate the way that the brain systems do. Some of those operating principles are just revealed to you with the negative correlation, and particularly with the pattern.

In terms of biological computers that are based on neural network principles, well clearly we're beginning at least to reveal the processes by which neural networks are functioning, and they may ultimately allow us to understand at what point something is perceived consciously as opposed to unconsciously. But what we can do, and we're already working with physicists and chemists who have the technology to fabricate nano-scale organic systems using molecular electronics, you can actually fabricate something that's like a brain network. The advantage is it's nano-scale – it's at least a thousand times smaller than your average brain circuit, and this is just some of what we're trying to do with these physicists and chemists in other parts of Europe. This is the traditional technology, the one we know about, in terms of nerve cells and axons, and we are constructing artificial versions of them using molecular electronics and chemistry. Not only can you produce something that's a thousandth of the size, but it can also transmit information bi-directionally, not like the nervous system, so in both directions, and probably hundreds if not thousands of times faster – whatever level you look at. What we're doing is creating these systems, or at least the physicists and the chemists are, giving them the algorithms to make them function like a brain, because that's the secret. They can do what they like in creating the system, but unless they actually use the mathematical algorithms that we derive from the way the brain does things to form this system so it functions like a brain, it will not be able to mimic the aspects of what we expect brains to be able to do.

You will recall Marvin from the Hitchhiker's Guide had too much of a brain. Artificial brains in the future will be undoubtedly extraordinarily small. We already have that technology. They will also be extraordinarily fast. My contention is that it's just a matter of scaling up from these systems. It's an inherent property of brain circuits that they should be capable of conscious experience, and so if we scale up, they should be conscious as well.

And yes – there's nothing to stop us, not just having the problem of getting rid of the keyboard, but also improving our own brains so that we can do even more than we do already!